

AD-A154 554

REVIEW AND ANALYSIS OF PHASE I INSTALLATION RESTORATION 1/1

PROGRAM REPORTS F. (U) HAZARDOUS MATERIALS TECHNICAL

CENTER ROCKVILLE MD G O PETERS FEB 85 DEVP-85-001

UNCLASSIFIED

DLA900-82-C-4426

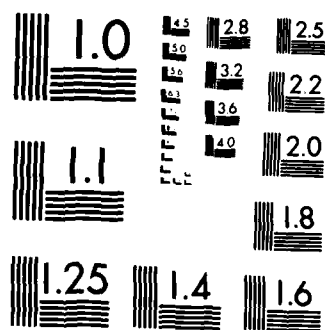
F/G 13/2

NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

1

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER <b>DEVP 85-001</b>	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Review and Analysis of Phase I Installation Restoration Program Reports for Selected Air Force Facilities		5. TYPE OF REPORT & PERIOD COVERED Final
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Gerald O. Peters		8. CONTRACT OR GRANT NUMBER(s) DLA900-82-C-4426
PERFORMING ORGANIZATION NAME AND ADDRESS Hazardous Materials Technical Center, The Dynamac Building, 11140 Rockville Pike, Rockville, MD 20852		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 111062058
CONTROLLING OFFICE NAME AND ADDRESS HQ AFESC/DEVP Tyndall AFB FL 32403		12. REPORT DATE <b>FEBRUARY</b> January 1985
MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		13. NUMBER OF PAGES 62
		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution is unlimited. Copies of this document can be ordered from the U.S. Department of Commerce, National Technical Information Service, Springfield, Virginia 22161; telephone: 703-487-4650.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES Project Officer, B. Lindenberg; tel.: 904-283-6193		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Installation Restoration Program (IRP) Site ranking model Multiple Regression Equation Hazard Assessment Rating Methodology (HARM) Hazardous Materials Uncontrolled hazardous waste disposal sites		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  See Reverse side		

DTIC  
ELECTE  
MAY 29 1985  
B

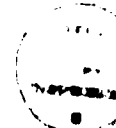
DTIC FILE COPY

AD-A154 554

A model was developed that ranks U.S. Air Force installations on the basis of potential for contamination from uncontrolled hazardous material disposal sites. The model is a multiple regression equation formulated from data presented in 77 records search reports for previously studied Air Force installations. Data required for the equation should be readily available from military archives and published soil survey reports. The model can be used to assign priorities to open or closed Air Force installations for initiating Phase I of the Installation Restoration Program.

The effects of statistical variance in the multiple regression equation are graphically displayed so that program managers can balance program goals and installation investigation costs.

Accession For	
NTIS STAG	<input checked="" type="checkbox"/>
ETIC CAN	<input type="checkbox"/>
Usama's	<input type="checkbox"/>
Just	<input type="checkbox"/>
PER CALL JC	
A-1	



REVIEW AND ANALYSIS OF PHASE I  
INSTALLATION RESTORATION PROGRAM  
REPORTS FOR SELECTED AIR FORCE FACILITIES

February 1985

Prepared for

Air Force Engineering and Services Center  
Tyndall Air Force Base, Florida 32403

Prepared by

Hazardous Materials Technical Center  
The Dynamac Building  
11140 Rockville Pike  
Rockville, Maryland 20852

Approved for public release; distribution is unlimited.  
Copies of this document can be ordered from the U.S.  
Department of Commerce, National Technical Information  
Service, Springfield, Virginia 22161; telephone:  
703-487-4650.

## Table of Contents

	<u>Page</u>
Executive Summary . . . . .	vi
Recommendations . . . . .	vii
Introduction . . . . .	1
Objective . . . . .	2
Approach . . . . .	2
Report Review . . . . .	4
Data Evaluation . . . . .	4
Statistical Analysis . . . . .	8
Uses of The Multiple Regression Equation . . . . .	14
Glossary . . . . .	21
Bibliography . . . . .	25
Appendix A - IRP List . . . . .	A-1
Appendix B - Soil Permeability Chart . . . . .	B-1
Appendix C - Data . . . . .	C-1
Appendix D - HARMSUM Histogram . . . . .	D-1

# Table of Figures and Tables

	<u>Page</u>
Figure 1. Data Summary Sheet . . . . .	6
Figure 2. Histogram of HRDIFF . . . . .	17
Figure 3. Certainty Curves . . . . .	19
Table 1. Correlation Coefficients for Pairs of Dependent and Independent Variables . . . . .	10
Table 2. Coefficients of Determination for Multiple Regression Equations . . . . .	13
Table 3. Actual and Predicted HARMSUMs Ranks for 78 Installations . . . . .	15,16

## EXECUTIVE SUMMARY

The Hazardous Materials Technical Center was retained by the Air Force Engineering and Services Center (AFESC) to develop a model for ranking Air Force Bases according to the contamination potential represented by existing hazardous material disposal sites. It is intended that the model be used to assign priorities to Air Force Bases for initiation of Phase I, of the Installation Restoration Program.

The model described here is a multiple regression equation. The equation's variables were selected for inclusion on the basis of both qualitative and quantitative factors. Numerical values in the equation were determined by multiple regression analysis of data reported in 78 Phase I Installation Restoration Program reports previously prepared for Air Force installations. As an index to contamination potential, the multiple regression equation uses the total of all hazard rating scores for hazardous material disposal sites discovered on an installation. Variables in the equation used to predict this index are the number of landfills, numbers of years that landfills were in use, number of fire training areas, and a semiquantitative rating of soil permeability. These four variables are expected to be available even for closed installations.

The variance between actual values of the index variable for the 78 Phase I installations and values predicted by the equation is such that comparisons between individual Air Force installations on the basis of predicted values will be statistically valid only where one value is much higher than another. However, the equation is still useful for ranking installations. It must be understood that some installations will be selected for study that turn out not to be of the highest priority, and others that should be studied may be delayed. A graphical technique is presented to assist program managers balancing such selection errors against the costs of preparing Phase I reports.

## Recommendations

Development of the installation ranking model was based in part on key assumptions regarding availability of data and statistical similarity between studied and unstudied installations. The model cannot be considered completely developed until these assumptions are verified. The following steps are recommended to accomplish this:

1. Locate and review archival information and published soil studies for 40 to 50 closed Air Force installations.
2. Tabulate data for the model's four independent variables: number of landfills, number of years that any landfill was in active use on the installation, number of fire training areas, and soil permeability or soil type. Soil data will have to be converted to the semiquantitative scale presented in Appendix B.
3. Also tabulate archival data on variables which may be incorporated in the model but were excluded because of the assumption that the data was not available. Number of tenants generating hazardous wastes, for instance, if it turns out to be available could strengthen the model.
4. During the review of archival information note any extenuating circumstances that suggest the need for Phase I study above and beyond the independent variables. Documented groundwater contamination, and explosions or other waste-related accidents at disposal sites are examples of such extenuating circumstances.
5. Examine data on the independent variables for completeness. Evaluate the effects of data deficiencies on the utility of the model.
6. Statistically analyze data distribution of the independent variables. For each variable, test the hypothesis that the unstudied installations are similar to the studied installations.
7. Using the model, calculate predicted HARMSUM's for each unstudied installation for which data is available on all four independent variables. Statistically test the hypothesis that the distribution of these predicted HARMSUM's are similar to the actual HARMSUMs of the previously studied installations.
8. Document the conclusions of steps 1-7.
9. If the assumptions regarding data availability and similarities between previously studied and unstudied installations hold up, complete the archival data retrieval for all closed Air Force bases. Calculate HARMSUMs and rank the installations.

## INTRODUCTION

In 1975, the Department of Defense (DOD) began its Installation Restoration Program (IRP) to assess past activities on DOD installations related to storage and disposal of hazardous materials and wastes. DOD policy is to identify and fully evaluate suspected problems associated with hazardous material disposal sites and to control hazards to health and welfare that may have resulted from these past activities. This policy is implemented through the four-phase approach of the IRP.

The Records Search comprises Phase I of the IRP and is intended to identify possible hazardous waste contaminated sites and to assess the potential for contamination migration beyond the installation boundaries. Phase II consists of field work and laboratory analysis to confirm the presence or absence of environmental contamination. Development of new technology is performed in Phase III. Phase IV consists of feasibility analysis, and design and construction of remedial measures that control the identified hazards.

At this writing (December, 1984), the U.S. Air Force has entered all of its operating installations in the IRP. The Phase I Records Search is under way or completed for these installations. Phase I has been completed for a substantial number of operating installations and a portion of these are progressing through Phase II.

However, few closed Air Force installations have been initiated into the IRP. The conclusiveness of Phase I studies depends in part on the oral reports of waste management personnel. These personnel will not be available in most cases for closed installations, so surveillance and sampling costs in Phase II are expected to be substantially higher than for operating installations. Stated another way, the closed installations are competing for available funding under the handicaps of poor initial data availability and generally higher expected costs for site identification and confirmation.

Nevertheless, the existence on closed installations of some disposal sites requiring remediation is highly likely. A method that would identify installations with the highest probability of containing problem sites would allow the Air Force to start the IRP for the worst installations first and to preserve as much available funding as possible for continuation of IRP activities at installations already in the program.

#### OBJECTIVE

The objective of this study is to develop a simple model that can be used to identify closed Air Force installations that are most likely to contain high priority hazardous waste disposal sites.

#### APPROACH

Completed Phase I reports have documented a substantial amount of data on hazardous waste disposal sites, past waste generation and disposal practices, and environmental settings at the sites. All of the analyses and results reported here are based on that information. The approach used is, therefore, empirical as compared to theoretical waste disposal site ranking methods, which assume a series of semiquantitative associations between the severity of contamination and various factors such as amount and type of waste, hydrogeologic setting, and uses of affected resources.

The model developed during this study differs from most other waste disposal site ranking methods in two other important aspects: it is applicable to ranking of installations where more than one disposal site may be located (most models rank individual sites), and it is based on statistical analysis, not contaminant dispersion computations or multicomponent scoring systems.

The Phase I reports contained data on many more parameters than could be used in the model. The primary reason for rejecting parameters was unavailability of that information for closed bases. The final model incorporates only parameters that are expected to be readily available from either military archives, national atlases, or local soil maps.

The volume of Phase I data available enables the application of statistical techniques in development of the model. Specifically, multiple regression analysis was applied to relate several factors that contribute to the number and hazard of disposal sites at an installation to a single index of number and hazard of sites. The resulting regression equation can be stated as follows:

$$y = b_0 + b_1A + b_2B + b_3C + \dots$$

Where:

y = values of the index of number and hazard of disposal sites, the dependent variable

A, B, C = values of the factors for which data are readily available and that contribute to the number and hazard of disposal sites, the independent variables

$b_0, b_1, b_2 \dots$  = regression parameters, constants determined mathematically from the Phase IV report data

Using this equation, with regression parameters calculated from the Phase I report data, the available data (represented by A, B, C, etc.) from a closed base can be substituted into the equation and an estimate of the index, y, can be calculated for that base. The utility of this estimate of the index is, of course, subject to both the variability of the Phase I data and the unprovable assumption that the Air Force installations that have not been studied are statistically similar to the ones that have been studied. Utility of the estimate will be discussed after development of the equation.

Development of the regression equation is discussed in the following three sections: Phase I report review and data compilation; data evaluation; and statistical analysis.

## REPORT REVIEW

The 78 draft and final Phase I reports used in this study are identified in Appendix A. Samples of these reports were read to identify the types of data that would be relevant to this study and that are typically reported in the Phase I reports. The intent of this parameter identification step was to be inclusive so that all potentially useful types of data would be inspected. The identified information types were included on the data summary sheet, reproduced as Figure 1; this form was used to record data from each Phase I report.

Most of the Phase I reports completed after 1982 contained the information required to complete the data summary sheet. Earlier reports lacked a consistent format and did not report all of the data called for on the data summary sheets. Nevertheless, only 1 report out of the 78 lacked data on any of the 5 variables ultimately chosen to create the model.

## DATA EVALUATION

The next step in constructing the model was to screen out variables that, based on inspection of the tabulated data, were believed to be unsuitable. The variables that were excluded at this point and reasons for excluding them are:

- "Major Command," "Mission of Base and/or Type of Aircraft," "Base Activities," and "Waste Types" - Base missions of Tactical Air Command (TAC) and Strategic Air Command (SAC) installations did not differ significantly in regard to operations or activities that generate wastes. Each installation maintains its structural plant, vehicles, and planes with methods and materials similar to other installations. Basic operations at most TAC and SAC installations also include engine test cells, nondestructive inspection labs, photo developing, grounds maintenance, and an auto hobby shop. As a result of the similar activities and types of materials used, the types of wastes generated by TAC and SAC installations are also similar. 66 of the 78 Phase I

The range of the HRDIFF distribution is such that an installation that actually ranks at the mid-point (rank 39) has a little more than a 5% chance of being predicted as having either the highest or lowest ranked site. It can be shown statistically that the predicted ranks for any two of the 77 installations would have to be separated by 35 ranks to conclude (at a 90% confidence level) that one is actually higher than the other. Using the same statistical principles with the unranked HARM scores, we found that the 90% confidence interval of 1,177 points is more than half the range of points between the lowest and highest HARM scores for the 78 installations i.e., 2,158 points.

It is concluded from those statistics that the regression equation is too weak to use for comparing individual installations in regard to the selected index of contamination potential, HARMSUM.

However, the equation can be used as a program management tool to select groups of installations for initiation of the IRP. In this use, the variance between predicted and actual results is not reflected as errors in comparing one installation to another. The variance will result unavoidably in unnecessary costs for studying installations that do not have serious disposal problems, and in delay of investigations of some installations with real problems. The regression equation can be used to estimate both the unnecessary costs and the probability that problem installations are not studied in a timely manner. With these estimates in hand, program managers may make judgments as to how many installations should be studied. The regression equation, if first applied to all unstudied installations, will indicate which installations are best to study.

The statistics on HRDIFF discussed earlier were used to formulate Figure 3, which shows a set of "certainty curves."

Figure 3 allows program managers to estimate the proportion of unstudied installations that would have to be studied (vertical axis) given the portion of highest ranked installations expected to be identified (horizontal axis) and the certainty that installations in that portion will, in fact, be identified (diagonals).

Figure 2. Histogram of HRDIFF

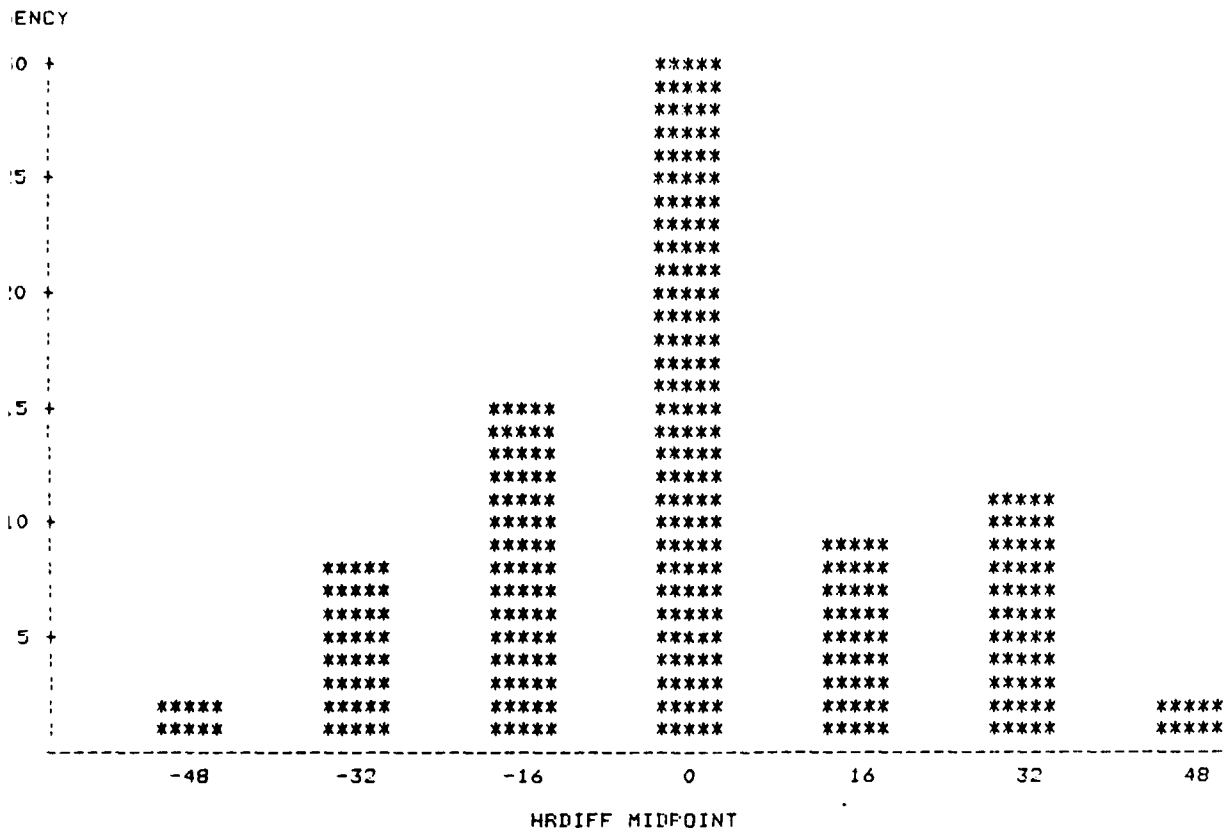


Table 3. Actual and Predicted HARMSUMs (Continued)

OBS	BASE	NUMBER	HRANK	HRDIFF
57	GRIFFISS	57	.	.
58	CANNON	58	61	-3
59	MYRTLE	59	63	-4
60	KIRTLAND	60	67	-7
61	ENGLAND	61	46	15
62	GEORGE	62	69	-7
63	MCGUIRE	63	62	1
64	LORING	64	14	50
65	MATHER	65	75	-10
66	NORTON	66	50	16
67	BERGSTROM	67	41	26
68	PLANT42	68	28	40
69	CHARLESTON	69	68	1
70	MARCH	70	70	0
71	HOLLOMAN	71	34	37
72	CASTLE	72	60	12
73	KELLY	73	59	14
74	WRIGHT-PAT	74	74	0
75	ELMENDORF	75	48	27
76	EILSON	76	43	33
77	MCCHORD	77	71	6
78	MCCLELLAN	78	77	1

Table 3. Actual and Predicted HARMSUMs

ORIS	BASE	NUMBER	HRANK	HRDIFF
1	BURLINGTON	1	6	-5
2	DESMOINES	2	2	0
3	PLANT78	3	3	0
4	KINGSLEY	4	33	-29
5	DOBBINS	5	15	-10
6	PLANT3	6	5	1
7	PLANT83	7	1	6
8	PLANTS2829	8	11	-3
9	LUKE	9	17	-8
10	VANCE	10	10	0
11	PLANT85	11	12	-1
12	TWINCITIES	12	4	8
13	HANCOCK	13	20	-7
14	COLUMBUS	14	73	-59
15	MCENTIRE	15	49	-34
16	WHEELER	16	55	-39
17	LOWRY	17	31	-14
18	CHANUTE	18	23	-5
19	RICHARDSB	19	36	-17
20	OLMSTED	20	53	-33
21	OTIS	21	24	-3
22	SHEPPARD	22	16	6
23	WILLIAMS	23	32	-9
24	MOODY	24	42	-18
25	SELFRIEDGE	25	38	-13
26	TYNDALL	26	57	-31
27	WESTOVER	27	35	-8
28	SEYMOUR	28	54	-26
29	REESE	29	76	-47
30	CLEAR	30	22	8
31	DEW	31	45	-14
32	DULUTH	32	18	14
33	EDWARDS	33	19	14
34	KEESLER	34	39	-5
35	MOUNTAINH	35	44	-9
36	MAXWELL	36	66	-30
37	WHITEMAN	37	27	10
38	O'HARE	38	8	30
39	SHAW	39	37	2
40	HILL	40	26	14
41	LANGLEY	41	72	-31
42	MACDILL	42	56	-14
43	TRAVIS	43	40	3
44	PLANT6	44	7	37
45	ROBINS	45	58	-13
46	NELLIS	46	65	-19
47	TINKER	47	47	0
48	HOMESTEAD	48	9	39
49	HANSCOM	49	29	20
50	NIAGRA FALL	50	13	37
51	CARSWELL	51	64	-13
52	PATRICK	52	51	1
53	FEASE	53	25	28
54	HICKAM	54	30	24
55	DAVISHON	55	52	3
56	BEALE	56	21	35

## USES OF THE MULTIPLE REGRESSION EQUATION

While the independent parameters used in the regression equation might be causally related to contamination potential, no such cause and effect relationships should be inferred from the statistical analysis reported here. Indeed, the HARM scores themselves are only semiquantitative estimates. And many factors that are included in calculating the scores during Phase I analysis are not included in the selected regression equation. The equation can only be used for the ranking of installations for further analysis - the purpose for which it was developed. Any applications that depend on any assumption of cause and effect between the dependent and independent variables are discouraged.

To illustrate the strength of the regression equation in ranking individual installations, both the actual HARMSUMS and the predicted HARMSUMS for 77 of the completed Phase I installations have been ranked and the 2 rankings compared. (One installation lacked data on LANDYR, so its predicted rank could not be calculated.) Table 3 lists the 78 installations in order of their actual HARMSUM. (Actual HARMSUM data for each installation is presented in Appendix C and a histogram of the actual HARMSUM scores is included in Appendix D.) Actual HARMSUM ranks are shown in the column labeled "NUMBER." The column "HRANK" shows the ranks for 77 installations predicted with the selected regression equation. Column "HRDIFF" shows the differences between actual and predicted rankings. The statistical distribution of HRDIFF values can then be used to describe the range and probabilities of predicted rankings given any assumed actual value.

Figure 2 is a plot of increments of HRDIFF values against the number of HRDIFF values in each increment. The information in Figure 2 also can be expressed by saying:

- o 39% of the predicted ranks are within 8 ranks of actual
- o 70% of the predicted ranks are within 24 ranks of actual
- o 95% of the predicted ranks are within 40 ranks of actual
- o 5% of the predicted ranks are more than 40 ranks from actual

Table 2. Coefficients of Determination for Multiple Regression Equations

Regression Equations with Number of Sites Recommended  
for Phase II as the Dependent Variable

<u>Number of Independent Variables</u>	<u>Coefficients of Determination</u>	<u>Variables in Equation</u>
2	0.12426526	LANDYR FIRE
2	0.13107985	LANDYR SOIL
2	0.13835324	FIRE SOIL
2	0.24149601	LANDFILL FIRE
2	0.26208397	LANDYR LANDFILL
2	0.28668836	LANDFILL SOIL
<hr/>		
3	0.17089895	LANDYR FIRE SOIL
3	0.26468588	LANDYR LANDFILL FIRE
3	0.29225529	LANDFILL FIRE SOIL
3	0.29798538	LANDYR LANDFILL SOIL
<hr/>		
4	0.30186771	LANDYR LANDFILL FIRE SOIL
<hr/>		

Regression Equations with HARMSUM  
as the Dependent Variable

<u>Number of Independent Variables</u>	<u>Coefficient of Determination</u>	<u>Variables in Equation</u>
2	0.14839593	LANDYR SOIL
2	0.21143562	FIRE SOIL
2	0.24595020	LANDYR FIRE
2	0.29145539	LANDFILL SOIL
2	0.31386748	LANDYR LANDFILL
2	0.31684753	LANDFILL FIRE
<hr/>		
3	0.26619339	LANDYR FIRE SOIL
3	0.32586046	LANDYR LANDFILL SOIL
3	0.34502282	LANDFILL FIRE SOIL
3	0.35543293	LANDYR LANDFILL FIRE
<hr/>		
4	0.37025993	LANDYR LANDFILL FIRE SOIL
<hr/>		

to random error. Table 2 lists each combination of dependent and independent variables and shows the  $R^2$  for each combination's regression equation.

Note in Table 2 that the equations that use the sum of an installation's HARM scores (HARMSUM) have higher coefficients of determination ( $R^2$ ) than those using the number of sites recommended for Phase II analysis (RECSITE). The inclusion of all four independent variables produces the highest  $R^2$  of any combination of independent variables. Therefore, for the group of Phase I reports used as the data base, the equation that yields the greatest strength for predicting an installation's potential for contamination is:

$$\begin{aligned}\text{HARMSUM} = & 360.4 + (7.684 \times \text{LANDYR}) \\ & + (34.25 \times \text{LANDFILL}) \\ & + (69.06 \times \text{FIRE}) \\ & - (59.04 \times \text{SOIL})\end{aligned}$$

where

HARMSUM = Total of HARM SUMs for all hazardous waste disposal sites on an installation

LANDFILL = Number of landfills on an installation

LANDYR = Number of years that any landfill was in active use on an installation

FIRE = Number of training pits

SOIL = A semiquantitative rating of soil permeability as described in Appendix B.

If the assumption is accepted that the installations used to develop this equation fairly represent the group of installations that have not been studied, then the equation can be used to predict our selected index of contamination potential, HARMSUM, for unstudied installation. The predicted HARMSUM values then can be used to rank the unstudied installations. Because data for the variables incorporated in the model are expected to be readily available, the ranking can be accomplished with a minimum of time and cost.

over the other site count parameters including the total of all identified sites, and "Landspill," the number of all sites except fire training pits and unspecified sites. These last variables, "Landspill" and total sites, include spills and other sites which are unlikely to be recorded in the archives. This difference in availability of the data is expected to more than compensate for the lower correlation coefficients associated with the "Landfill" and "Fire" parameters.

Quantitative and qualitative screening of the independent variables leaves four to be evaluated for inclusion in the multiple regression equation that is the basis of the installation ranking model. The four independent variables are:

- o Number of fire training pits (Fire)
- o Number of landfills (Landfill)
- o Number of years that landfills were in use (Land Year)
- o Soil permeability rating (Soil)

There are 11 combinations of these 4 independent variables: 6 using 2 variables, plus 4 combinations using 3, plus 1 with all 4. A regression equation can be developed for each of these combinations with each of the 2 dependent variables for a total of 22 possible equations.

To select the regression equation with the strongest predictive capability, the 22 equations were analyzed by STEPWISE and RSQUARE procedures available through Statistical Analysis System (SAS Institute, Inc., 1982). The RSQUARE procedure ranks combinations of independent variables regressed against the dependent variables using the coefficient of determination, also called the square of multiple-correlation coefficient and designated as  $R^2$ .  $R^2$  measures that part of the variation between actual and predicted values of the dependent variable that is due to the regression equation rather than

Table 1. Correlation Coefficients for Pairs of  
Dependent and Independent Variables

	<u>Total</u>	<u>Landspill</u>	<u>Landfill</u>	<u>Land Year</u>	<u>Fire</u>	<u>Fire Year</u>	<u>Soil</u>	<u>Ground</u>	<u>Rain</u>	<u>Tenant</u>	<u>Shop</u>
*Recsite	.529 .0001	.499 .0001	.481 .0001	.296 .0089	.257 .0228	.124 .3256	-.285 .0114	-.120 .3111	-.119 .3384	.262 .0211	.140 .2447
HARM SUM	.658 .0001	.723 .0001	.516 .0001	.363 .0012	.393 .0004	.140 .2657	-.220 .0530	.065 .5839	-.166 .1798	.302 .0076	.113 .3476

Note: For each pair, the upper numbers are the correlation coefficients; the lower numbers are the probabilities that the pair is not linearly correlated.

\* Number of sites used for Phase II analysis.

correlation would be demonstrated by measuring the length and weight of two-by-fours cut from the same wood on the same mill. The correlation coefficient for those two variables would be very close to 1. Another positive correlation, but one not as strong, would be demonstrated by measuring the heights and weights of a group of men. If two variables show a negative correlation, such as the number of Christmas presents purchased and your bank account balance, the correlation coefficient will be between 0 and -1.

As a correlation coefficient approaches zero, whether it is positive or negative, we begin to conclude that there is no correlation between the two variables being examined. But how small should we let the coefficient get before concluding that a pair of variables is not strongly correlated enough to support the model? We answer this question by statistically testing the hypothesis that the coefficient is equal to zero, or in other words, that there is no linear association between them. The results of this test will tell us the probability of error if we use the pair of variables as if they were linearly associated. In the development of the installation ranking mode, the maximum probability of error we will accept is five percent.

Table 1 shows both the correlation coefficient for each pair of variables (upper numbers) and the probability that each pair is not linearly associated (lower numbers).

Based on these results, four of the independent variables were not considered further for inclusion in the installation ranking model. These were net number of years that fire training pits were used, depth to groundwater, precipitation, and number of shops generating wastes.

With these results in hand, several variables were again evaluated qualitatively. The number of tenant shops, although showing a reasonably strong correlation with the index parameters, was rejected because of the expectation that this information would not be readily available in military archives for many closed installations. Availability of data is also the justification for preferring number of fire pits and number of landfills

This variable is not subject to different reporting approaches among the Phase I contractors as are the HARM score sums, and the number is obviously related to how many sites may ultimately be designated for remedial action. On the other hand, the sum of HARM scores should more accurately reflect the aggregate degree of hazard for disposal sites on an installation rather than just the number of sites that require more study. However, it must be noted that Phase I contractors differed in their decisions regarding which sites to rate. Some contractors rated only those sites they deemed deserving of Phase II study; other contractors rated all sites, regardless of the apparent hazard; and others took intermediate approaches.

#### STATISTICAL ANALYSIS

The statistical analysis was conducted in three steps:

1. Calculation of correlation coefficients for each possible combination of index variables paired with other (independent) variables
2. Multiple regression analysis of independent variables versus the index variables
3. Analysis of the strength of multiple regression equations

In the process of analyzing the data, it became apparent that some of the independent variables, even though they correlated well with index variables, were not useful in the final model. These variables were not considered for use in the final model because it was determined that archival records would not yield the relevant data for closed bases. Following are a detailed discussion and the results of three steps in the statistical analysis.

A correlation coefficient is a measure of the degree of linear association between two variables. Two variables that are positively correlated will have a correlation coefficient between 0 and 1. The closer the coefficient is to 1, the better will be the accuracy in predicting one variable if a value for the other is given. An example of a strong positive

- o Number of years that waste disposal sites were in use
- o Number of sites recommended for Phase II study.
- o Number of fire training areas.
- o Number of years that fire training areas were in use..
- o Soil permeability (rated on a semiquantitative scale of 1 to 5, see Appendix B).
- o Groundwater depth in feet. -
- o Net precipitation in inches (annual rainfall minus annual evapotranspiration).
- o Number of tenants generating hazardous wastes.
- o Number of shops generating hazardous wastes.

Initial research in this study focused on "number of sites recommended for Phase II study" as the index of contamination potential for each installation. Subsequent consideration of the statistical analysis (discussed in the next section) led to consideration of the Hazard Assessment Rating Methodology (HARM) scores as an alternate index. The sum of HARM scores for all disposal sites that were rated on each installation was, therefore, added to the list of installation parameters. Another addition was made in response to concern that data from closed installations may not be available for use in the model on a number of spill sites, leaks and other incidental or one-time disposal sites. The "Number of Landfills" was, therefore, tabulated as a separate parameter from "total number of waste disposal sites."

Data for all variables that were analyzed statistically are recorded by installation in Appendix C.

Four variables were initially considered as candidates to be the index of contamination potential (the dependent variable). Two of these, "Evidence of Groundwater Contamination" and "Types of Contaminants," were eliminated prior to the statistical analysis for reasons already discussed. Of the remaining two, preference was given initially to "Number of Sites Recommended for Phase II Study."

reports were for TAC or SAC installations. Wastes generated by Air Force plants and Air Logistics Centers differ from the TAC and SAC wastes because these installations are more involved in industrial fabrication and production.

However, the small number of reports on plants (7) and logistics centers (5) was expected to prevent any statistically valid finding of differences between the major commands.

- o "Waste Quantities" - Data on waste quantities were unavailable for most identified waste streams. "Number of Tenant Shops," and "Years of Disposal Site Operation" were accepted as the next best estimates of waste quantity. (Note that these variables were also subsequently excluded.)
- o "Size of Facility" - The data on size of facilities was expressed in terms of average and was of little use because the percentage of developed land on the bases could not be determined. Large bases generally had comparable levels of maintenance and other waste generating activities as bases with more restricted boundaries.
- o "Topography" - Topography does not provide a basis for distinguishing between Air Force installations. In line with their primary mission, they are all flat.
- o "Evidence of Groundwater Contamination" - Because of the scope of Phase I Records Searches, documentation of groundwater contamination was spotty. Indeed a major purpose of Phase II work is to provide this documentation. Information on groundwater contamination in the Phase I reports was typically speculative and not sufficiently substantive for incorporation in the model.
- o "Types of Contaminants" - This variable did not provide a basis for distinguishing between Air Force installations. Recommended monitoring programs for each installation did not vary much from a basic list of contaminants that included volatile organic compounds, total organic carbon, total organic halogens, heavy metals, pesticides, and drinking water standard parameters.

The remaining parameters that were judged potentially useful in forming the model include:

- o The total number of waste disposal sites including landfills; spill sites; burial pits; hardfills; leaks; and diked, drained or trenched areas.

Figure 1. Data Summary Sheet

INSTALLATION

SOURCES:

- Major command
- Mission of base and/or type of aircraft
- Base activities
- Number of tenant shops
- Years (dates) of operation
- Methods and dates of disposal
- Number of landfills, spills, pits, burial sites, hardfills, trenches, dikes, and leaks, and number of sequential years used or occurred
- Number of fire training areas and number of years used

WASTES:

- Types
- Quantities

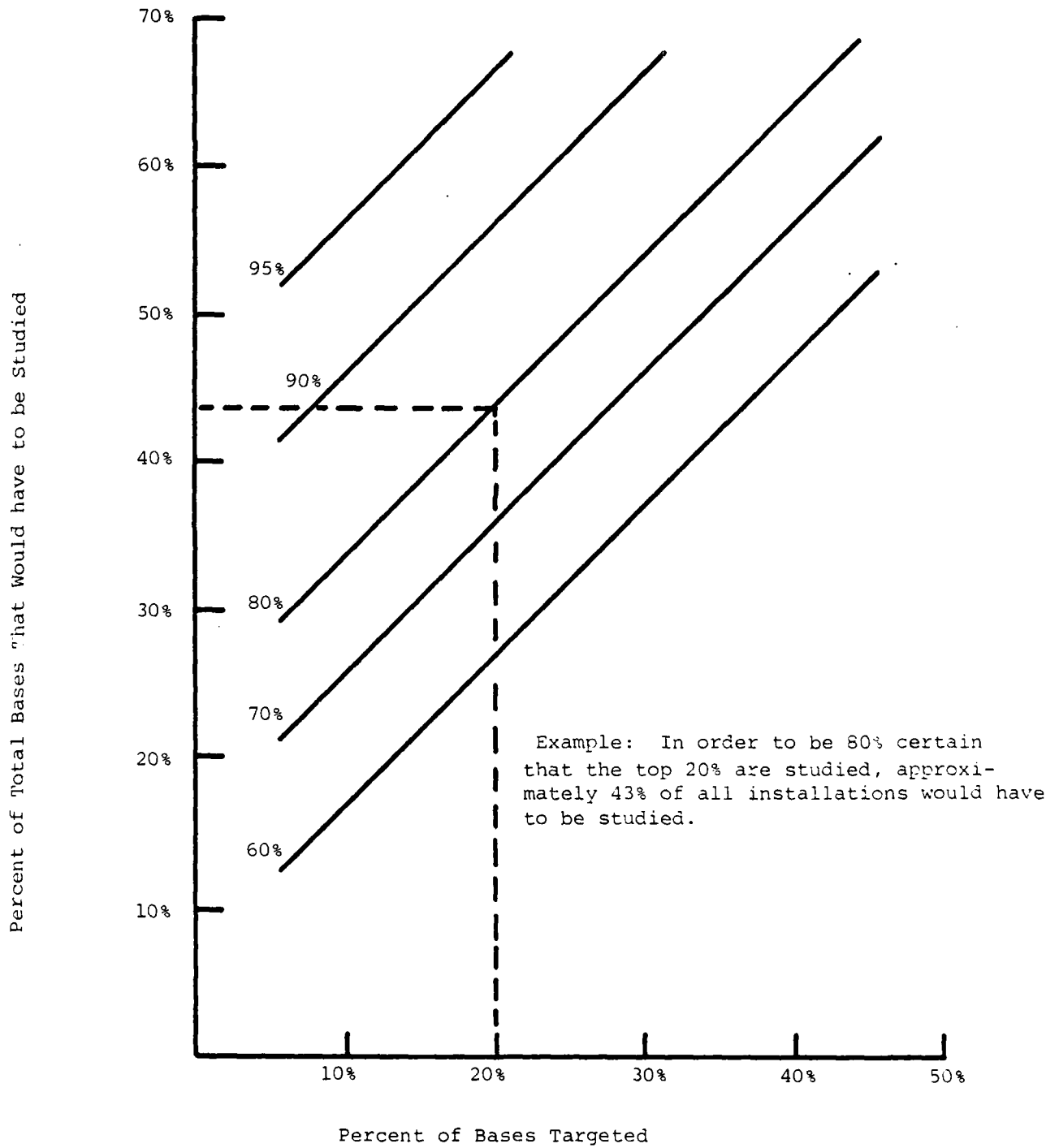
ENVIRONMENTAL FACTORS:

- Size of facility
- Depth to groundwater
- Soil types
- Precipitation
- Topography

RESULTS:

- Number of sites recommended for further study
- Evidence of groundwater contamination
- Types of contaminants

Figure 3. Certainty Curves



To illustrate, assume that a manager wants to identify the top 20% of a group of installations and that predicted HARMSUMs have been calculated for installations in the group using the regression equation. He would find the 20% point on the horizontal axis and draw a line vertically from it, as illustrated in the figure. The manager would then decide how much error he will allow in identifying the worst 20%. Assuming he has a restricted budget, he selects a low certainty level, say 80%. He then draws a horizontal line through the intersection of the 80% certainty line with the vertical line just drawn. The new line crosses the vertical axis at 43%. This tells him that, to meet his objectives, he should study the 43% of the installations with the highest predicted HARMSUMs.

Will he in fact identify each of the worst 20% of the installations? Maybe. But because he accepted an 80% certainty level, he would most likely identify four out of five of the worst 20%. He might find less or he might find all of the worst. (He will not know exactly what percent of the worst installations he actually identified until all of the sites are studied.)

Figure 3 can also be used in the other direction. For example, assume the program manager has a budget sufficient to study 30% of unstudied sites. If he draws a horizontal line through the 30% point on the vertical axis, he will be able to predict that he will identify the worst 6% with 80% certainty, the worst 15% with 70% certainty, and the worst 23% with 60% certainty.

Please note that Figure 3 does not predict HARMSUMs and does not rank installations - it only indicates what proportion of installations should be studied in light of the program manager's objectives. The utility of the certainty curves rest, as the regression equation does, on the assumption that the 77 previously studied installations fairly represent the unstudied installations.

## GLOSSARY

### Coefficient of Determination:

The fraction of the total variation in a dependent variable that is accounted for by the association between the dependent and independent variables.

### Confidence Interval:

A method of stating both how close the value of a statistic is likely to be to some specified value of a variable, and the probability of its being that close. The confidence interval is determined by an arbitrary degree of probability, the confidence level, appropriate to the problem at hand and by the variance of the variable.

### Dependent Variable:

For regression analysis, the variable that is to be predicted on the basis of one or more independent variables.

### "Fire":

Abbreviation of number of fire training areas.

### Fire Training Areas:

Locations at which fire training exercises were or are held. Usually consists of some structure or plane wreckage that is doused with flammable liquids and ignited, to be extinguished by trainees.

### HARM:

Hazard Assessment Rating Methodology, a multicomponent scoring system used by the Air Force to rate the contamination potential of individual hazardous waste disposal sites.

### HARMSUM:

The sum of HARM scores of all hazardous waste disposal sites identified on an installation during Phase I evaluations.

**Hardfill:**

Landfills or surface dumps consisting of construction rubble, abandoned equipment and other solids, and nonbiodegradable waste materials.

**Histogram:**

A picture of a number of measurements or observations. Measurements are grouped by selected intervals, shown normally on the horizontal axis of the histogram; and the number of measurements is demonstrated by the vertical height of a bar for each interval.

**Independent Variable:**

In regression analysis a variable that by itself or in combination with other independent variables will be used to predict the values of a dependent variable.

**"Landfill":**

Term used for number of sanitary landfills reported for an installation.

**"Landspill":**

Term used for number of waste disposal sites reported for an installation excepting fire training areas and some miscellaneous types of sites such as lagoons.

**"Landyear":**

Term used for number of years during which any landfills were in active use on an installation.

**Multicomponent Scoring System:**

Methods for ranking disposal sites that are based on semiquantitative scoring of a number of factors. The factors are included for their apparent connections with the potential for resource contamination and for the effects of such contamination upon resource use. The factors are grouped within two or more components, such as "waste characteristics," "transport route," and "resource use."

#### Multiple Regression Analysis:

A method for describing the joint relationship of a single variable, the dependent variable, to several independent variables. The analysis calculates a set of positive or negative coefficients which, when multiplied by the independent variables, and the products added together, yield an equation that predicts the dependent variable with a minimum of error.

#### Regression Equation:

An equation produced by single or multiple regression analysis.

#### RSQUARE:

A statistical analysis procedure available as a program in the Statistical Analysis System. The RSQUARE procedure performs all possible regressions for one or more dependent variables and a collection of independent variables, then reports the coefficient of determination,  $R^2$ , for each regression.

#### Semiquantitative:

Numbering schemes used to describe judgments or estimates, or to reduce quantitative measurements to common scales.

#### "Soil":

Abbreviation of soil permeability scored by the table in Appendix B.

#### Soil Permeability:

The capacity of a soil to transmit a fluid, typically water, under stated conditions of saturation, temperature, and hydraulic head.

#### STEPWISE:

A statistical analysis procedure available as a program in the Statistical Analysis System. STEPWISE provides a choice of five selection strategies for identifying regression equations that best fit the variables and data being evaluated. The selection strategy used in this report is called Maximum  $R^2$  Improvement

STEPWISE (Continued):

(MAXR). MAXR begins by finding the one-variable regression equation that has the highest coefficient of determination,  $R^2$ . Additional variables are added or substituted on the basis of  $R^2$  improvement until the best one-variable, two-variable, three-variable, etc. regression equations have been identified.

## Bibliography

1. SAS User's Guide Statistics. 1892 Edition; SAS Institute Inc.,  
Box 8000. Cary, North Carolina.

APPENDIX A

IRP LIST

# IRP LIST

1. Air Force Plant Number 3      Tulsa, Oklahoma  
Final  
December 1983
2. Air Force Plant Number 6      Marietta, Georgia  
Final  
March 1984
3. Air Force Plants Numbers 28 and 29      Everett and Lynn, Massachusetts  
Draft  
April 1984
4. Air Force Plant Number 42      Palmdale, California  
Final  
October 1983
5. Air Force Plant Number 78      Brigham City, Utah  
Draft  
January 1984
6. Air Force Plant Number 83      Albuquerque, New Mexico  
Draft  
November 1983
7. Air Force Plant Number 85      Franklin County, Ohio  
Final  
February 1984
8. Alaska DEW Line Stations  
Final  
October 1981
9. Beale Air Force Base      Marysville, California  
Final  
April 1984
10. Bergstrom Air Force Base      Austin, Texas  
Final  
July 1983
11. Burlington International Airport      Chittenden County, Vermont  
Final  
September 1983
12. Cannon Air Force Base      Clovis, New Mexico  
Final  
August 1983

13. Carswell Air Force Base Fort Worth, Texas  
Final  
February 1984
14. Castle Air Force Base Atwater, California  
Final  
October 1983
15. Chanute Air Force Base Rantoul, Illinois  
Final  
December 1983
16. Charleston Air Force Base Charleston County, South Carolina  
Final  
October 1983
17. Clear Air Force Station Anderson, Alaska  
Draft  
September 1981
18. Columbus Air Force Base Columbus, Mississippi  
Draft  
May 1984
19. Davis-Monthan Air Force Base Tucson Arizona  
Final  
August 1982
20. Des Moines Municipal Airport Polk County, Iowa  
Final  
September 1983
21. Dobbins Air Force Base Marietta, Georgia  
Final  
April 1982
22. Duluth International Airport Duluth, Minnesota  
Final  
March 1982
23. Edwards Air Force Base Rosamond, California  
Final  
April 1981
24. Eielson Air Force Base Fairbanks, Alaska  
Final  
November 1982
25. Elmendorf Air Force Base Anchorage, Alaska  
Final  
September 1983

26. England Air Force Base      Alexandria, Louisiana  
Final  
May 1983
27. George Air Force Base      Victorville, California  
Final  
January 1982
28. Griffiss Air Force Base      Rome, New York  
Final  
July 1981
29. Hancock Field      Syracuse, New York  
Final  
July 1982
30. Hanscom Air Force Base      Belford, Massachusetts  
Draft  
April 1984
31. Hickam Air Force Base      Oahu Island, Hawaii  
Final  
July 1983
32. Hill Air Force Base      Ogden, Utah  
Draft  
November 1981
33. Holloman Air Force Base      Alamogordo, New Mexico  
Final  
August 1983
34. Homestead Air Force Base      Homestead, Florida  
Final  
August 1983
35. Keesler Air Force Base      Biloxi, Mississippi  
Final  
April 1984
36. Kelly Air Force Base      San Antonio, Texas  
Draft  
November 1981
37. Kingsley Field      Klamath Falls, Oregon  
Final  
February 1982
38. Kirtland Air Force Base      Albuquerque, New Mexico  
Draft  
November 1981

39. Langley Air Force Base Hampton, Virginia  
Final  
June 1981
40. Loring Air Force Base Limestone, Maine  
Final  
January 1984
41. Lowry Air Force Base Denver, Colorado  
Final  
August 1983
42. Luke Air Force Base Glendale, Arizona  
Final  
February 1982
43. MacDill Air Force Base Tampa, Florida  
Final  
November 1981
44. March Air Force Base Riverside, California  
Final  
April 1984
45. Mather Air Force Base Sacramento, California  
Final  
June 1982
46. Maxwell Air Force Base Montgomery, Alabama  
Final  
January 1984
47. McChord Air Force Base Tacoma, Washington  
Final  
August 1982
48. McClellan Air Force Base Sacramento, California  
Final  
July 1981
49. McEntire Air National Guard Base Eastover, South Carolina  
Final  
January 1984
50. McGuire Air Force Base Wrightstown, New Jersey  
Final  
November 1982
51. Moody Air Force Base Valdosta, Georgia  
Final  
February 1983

52. Mountain Home Air Force Base      Mountain Home, Indiana  
Draft  
April 1983
53. Myrtle Beach Air Force Base      Myrtle Beach, South Carolina  
Final  
October 1981
54. Nellis Air Force Base      Las Vegas, Nevada  
Final  
April 1982
55. Niagra Falls Air Force Reserve Facility      Niagra Falls, New York  
Final  
December 1983
56. Norton Air Force Base      San Bernadino, California  
Final  
October 1982
57. O'Hare Air Reserves Forces Facility      Chicago, Illinois  
Final  
December 1983
58. Olmstead Air Force Base      Middletown, Pennsylvania  
(Harrisburg International Airport)  
Final  
April 1984
59. Otis Air National Guard Base      Falmouth, Massachusetts  
Final  
January 1983
60. Patrick Air Force Base      Cocoa Beach, Florida  
(Eastern Space and Missile Center)  
Draft  
April 1984
61. Pease Air Force Base      Portsmouth, New Hampshire  
Draft  
November 1983
62. Reese Air Force Base      Lubbock, Texas  
Draft  
April 1984
63. Richards-Gebaur Air Force Base      Belton, Missouri  
Final  
March 1983
64. Robins Air Force Base      Warner Robins, Georgia  
Final  
April 1982

65. Selfridge Air National Guard Base     Mt. Clemens, Michigan  
Final  
April 1983
66. Seymour-Johnson Air Force Base     Goldsboro, North Carolina  
Final  
July 1982
67. Shaw Air Force Base     Sumter, South Carolina  
Final  
May 1983
68. Sheppard Air Force Base     Wichita Falls, Texas  
Final  
February 1984
69. Tinker Air Force Base     Oklahoma City, Oklahoma  
Final  
April 1982
70. Travis Air Force Base     Fairfield, California  
Final  
August 1983
71. Twin Cities Air Force Reserve Base     Minneapolis, Minnesota  
Final  
March 1983
72. Tyndall Air Force Base     Panama City, Florida  
Final  
December 1981
73. Vance Air Force Base     Enid, Oklahoma  
Draft  
May 1984
74. Westover Air Force Base     Chicopee, Massachusetts  
Final  
April 1982
75. Wheeler Air Force Base     Waipahu, Hawaii  
Final  
July 1983
76. Whiteman Air Force Base     Sedalia, Missouri  
Final  
March 1984
77. Williams Air Force Base     Chandler, Arizona  
Final  
February 1984
78. Wright-Patterson Air Force Base     Dayton, Ohio  
Draft  
December 1981

APPENDIX B  
SOIL PERMEABILITY CHART

# Soil Permeability Chart

Soil	Permeability Mean in/hr*	Hydrologic Group
Gravel	1000	1.0
	550	1.5
Sand	100	2.0
	51	2.5
Silty Sand	2.34	3.0
	1.20	3.5
Silt, Loess	.025	4.0
	.005	4.5
Glacial Till, Clay	.002	5.0

- o Group 1.0-2.5 Soils having a high infiltration rate (low runoff potential) when thoroughly wet. These consist chiefly of deep, well-drained to excessively drained sands or gravels. These soils have a high rate of water transmission.
- o Group 2.5-3.5 Soils having a moderate infiltration rate when thoroughly wet. These consist chiefly of moderately deep or deep, moderately well-drained or well-drained soils that have moderately fine texture to moderately coarse texture. These soils have a moderate rate of water transmission.
- o Group 3.5-4.5 Soils having a slow infiltration rate when thoroughly wet. These consist chiefly of soils that have a layer that impedes the downward movement of water or soils that have moderately fine texture or fine texture. These soils have a slow rate of water transmission.
- o Group 4.5-5.0 Soils having a very slow infiltration rate (high runoff potential) when thoroughly wet. These consist chiefly of clay soils that have a permanent high water table, soils that have a claypan or clay layer at or near the surface, and soils that are shallow over nearly impervious material. These soils have a very slow rate of water transmission.

\* Source: Cherry, John A. and Freeze, R. Allen. Groundwater. Prentice Hall, Inc., Englewood Cliffs, New Jersey 07632. 1979. Page 29.

APPENDIX C  
DATA

## DATA

BASE	TOTAL	RECSITE	LANDSPIL	LANDYR	FIRE	FIREYR	SOIL	GROUND	RATN	TENANT	SHOP	LANDFILL	HARMSUM
BEALE	16	6	15	33	2	26	4.0	80.0	45.0	17	42	3	890
BERGSTROM	26	5	25	40	1	40	4.0	1500.0	-15.0	7	14	7	1203
BURLINGTON	2	2	1	20	1	20	4.0	80.0	9.0	1	9	2	130
CANNON	19	4	14	35	4	25	3.0	200.0	-55.0	3	11	5	938
CARSWELL	17	10	15	40	2	40	3.0	115.0	-25.0	5	14	9	796
CASTLE	38	28	35	40	3	36	2.5	40.0	-41.0	12	43	5	1367
CHANUTE	12	6	6	31	2	33	4.0	15.0	4.5	3	29	4	391
CHARLESTON	40	18	26	31	6	24	2.5	10.0	8.0	11	36	4	1330
CLEAR	16	8	15	30	1	1	3.0	70.0	.	1	.	4	558
COLUMBUS	17	2	13	41	4	12	4.0	10.0	15.0	7	40	13	351
DAVISHON	34	7	30	44	4	44	3.5	300.0	-55.0	6	22	2	836
DESMOINES	3	2	1	1	2	18	4.0	350.0	.	1	16	0	138
DEW	44	13	44	30	.	.	4.0	.	6.0	7	.	13	573
DORRINS	11	3	9	33	2	33	4.5	.	10.0	5	30	2	264
DULUTH	13	5	8	25	2	22	4.5	12.0	10.0	5	17	5	574
EDWARDS	12	7	12	25	1	.	2.5	65.0	.	1	.	4	588
EILSON	42	17	27	32	2	7	3.0	5.0	0.0	3	21	6	2026
ELMENDORF	54	13	33	46	1	44	3.0	20.0	.	12	41	6	1542
ENGLAND	38	6	34	30	4	41	4.5	5.0	8.0	7	23	6	988
GEORGE	54	14	54	40	.	.	1.5	100.0	-80.0	15	31	13	1033
GRIFFISS	12	4	7	.	.	.	4.0	67.0	.	.	.	7	892
HANCOCK	9	4	5	29	1	55	3.5	5.0	10.0	5	17	5	347
HANSCOM	22	12	19	35	2	23	3.0	3.0	.	7	49	2	786
HICKAM	22	6	20	25	2	43	4.0	1.5	-60.0	12	83	6	831
HILL	15	4	13	28	2	25	4.0	420.0	13.0	5	14	5	702
HOLLOMAN	43	1	42	37	1	37	3.0	7.5	-59.0	6	15	4	1363
HOMESTEAD	18	8	15	15	3	34	4.5	7.0	64.0	14	60	1	760
NEESLER	20	9	15	38	1	29	2.0	6.0	13.0	11	23	3	611
KELLY	25	14	22	44	2	50	4.0	13.0	-30.0	6	22	8	1498
KINGSLEY	12	3	7	27	2	17	3.0	5.0	14.0	1	.	4	205
KIRTLAND	17	5	15	32	2	.	1.5	100.0	-22.0	7	35	10	978
LANGLEY	22	8	15	34	3	30	3.0	50.0	-5.0	33	52	10	716
LORING	19	11	18	35	1	32	5.0	10.0	31.0	2	13	4	1058
LOWRY	23	2	7	39	2	19	3.5	10.0	-30.0	6	28	2	382
LUKE	16	1	14	26	2	27	4.5	300.0	-53.0	14	47	4	313
MACDILL	23	9	22	43	1	28	2.0	3.0	-6.0	5	17	6	717
MAHCH	30	14	15	44	3	44	2.0	13.0	-70.0	11	41	7	1350
MATHER	19	6	19	54	4	54	1.5	100.0	-27.0	11	28	7	1084
MAXWELL	73	10	10	53	3	70	3.5	5.0	-6.0	10	32	6	618
MCCHORD	54	10	48	45	6	30	3.5	100.0	20.0	9	44	6	2146
MCCLELLAN	45	24	38	40	7	40	3.5	100.0	-26.0	10	10	35	2288
MCCENTIRE	13	4	9	39	4	39	2.5	37.0	20.0	1	11	1	359
MCGUIRE	29	10	19	36	3	38	2.5	50.0	9.0	9	28	6	1054
MOODY	14	4	12	37	2	18	3.0	15.0	5.0	1	12	4	466
MOUNTAIN	17	5	12	28	5	39	4.0	400.0	-27.0	5	27	3	613
MURLE	15	10	11	28	4	28	2.0	3.0	.	9	19	5	959
NETTIS	33	4	30	42	3	40	2.5	50.0	-68.0	17	81	6	750
NIAGARAFALL	20	9	11	25	3	28	4.5	4.0	8.6	2	30	1	791
NORTON	20	12	17	40	3	40	2.0	60.0	-41.0	9	22	2	1129
O'HARE	16	9	14	30	1	30	5.0	15.0	4.0	3	18	2	631
OLMSTED	10	6	7	32	3	32	2.5	30.0	.	14	36	5	431
OTTIS	6	3	2	33	2	33	2.5	50.0	21.0	1	.	1	434
PATRICK	26	14	18	42	2	33	2.0	100.0	.	6	30	6	813
PEASE	18	1	16	30	2	28	5.0	20.0	19.0	6	20	6	814
PLANTS2829	24	3	8	34	1	.	3.0	3.5	18.0	1	32	0	307
PLANTS	12	3	7	21	1	33	4.0	.	-14.0	2	27	0	270

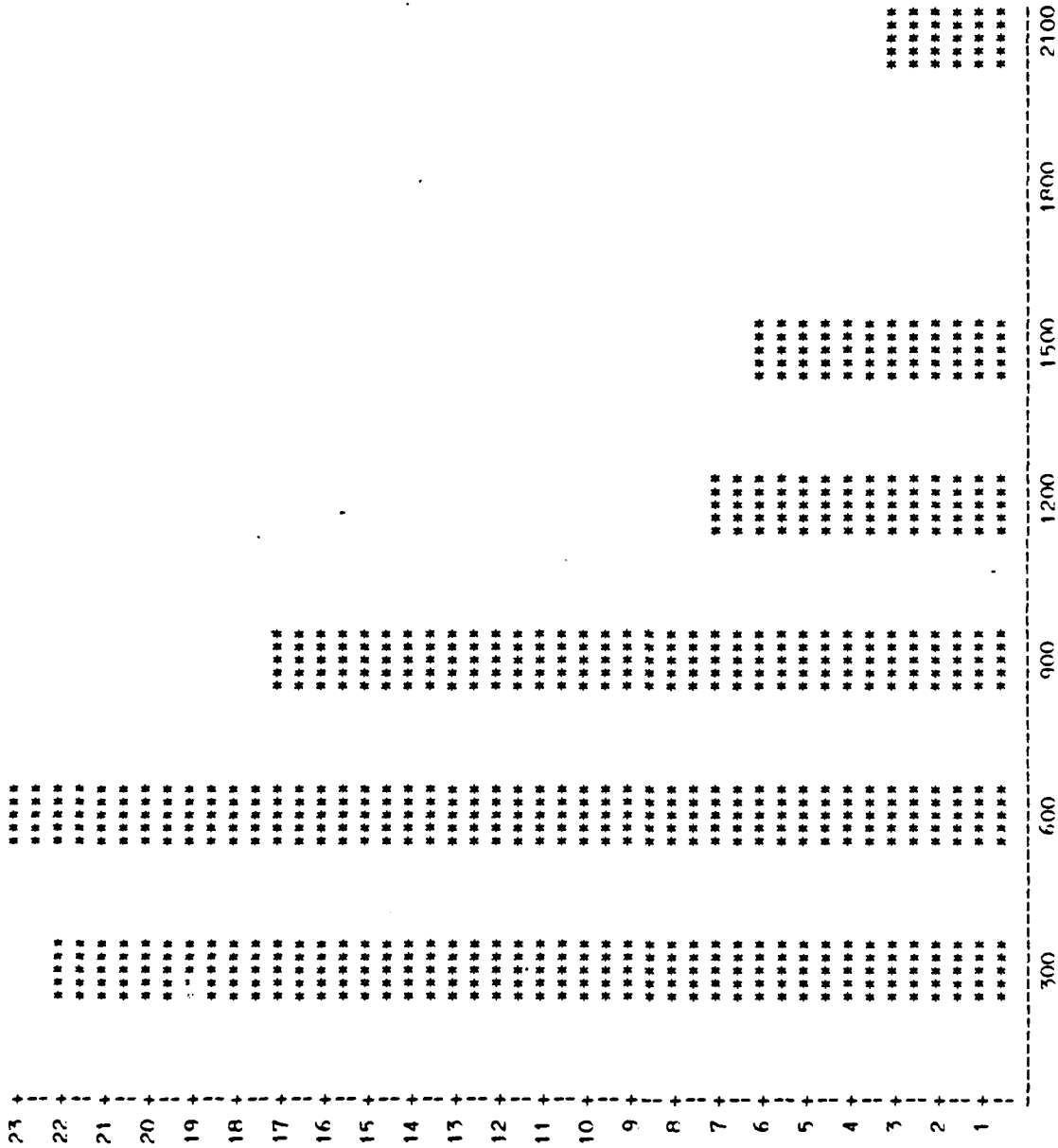
## DATA

BASE	TOTAL	RECSITE	LANDSPIL	LANDYR	FIRE	FIREYR	SOIL	GROUND	RAIN	TENANT	SHOP	LANDFILL	HARMSUM
PLANT42	31	7	28	30	3	30	3.0	300	-65	8	18	1	1283
PLANT6	12	10	12	35	.	.	4.5	20	13	2	30	2	737
PLANT78	5	4	4	20	.	.	4.0	145	-26	1	19	0	203
PLANT83	13	4	7	10	.	.	4.0	.	-54	1	17	0	291
PLANT85	9	4	7	42	2	37	5.0	25	4	1	12	0	331
REESE	37	10	20	33	6	.	5.0	60	.	5	23	14	546
RICHARDSO	9	2	5	23	4	28	4.0	.	-5	4	15	3	395
ROBINS	10	8	12	32	4	41	2.5	200	2	10	33	4	738
SELFRIDGE	7	7	5	48	2	30	4.5	12	0	1	.	3	469
SEYHOUR	9	5	8	33	3	28	2.0	6	9	9	23	4	525
SHAW	17	3	12	40	3	40	5.0	25	3	13	31	3	651
SHEFFARD	23	4	12	17	3	28	4.0	10	-37	11	42	3	437
TINKER	24	20	12	37	2	34	3.0	100	-27	7	25	6	758
TRAVIS	15	7	11	34	4	41	4.5	10	-31	14	45	3	727
TWINCITIES	8	1	7	11	.	.	3.0	10	-4	3	23	2	332
TYRIMALL	17	3	13	38	2	38	2.5	5	5	2	13	6	493
VANCE	8	8	7	20	1	30	3.0	2	.	5	23	3	318
WESTOVER	15	1	12	35	3	35	5.0	10	12	1	9	4	512
WHEELER	8	7	6	60	1	30	2.5	30	-24	10	20	3	367
WHITFMAN	21	6	13	37	1	.	4.0	150	-3	9	27	5	624
WILLIAMS	10	6	8	44	2	44	3.5	400	-65	8	29	1	464
WRIGHT-PAT	28	21	19	40	4	30	3.0	15	5	11	42	12	1518

APPENDIX D  
HARMSUM HISTOGRAPH

# HARMSUM Histogram

Number of Installations



Midpoint of HARMSUM Intervals

**END**

**FILMED**

**7-85**

**DTIC**